Appendix J – Common Record Physical Layout

Introduction

This section provides standards for the physical layout of the Common Record XML Document file. Standards on the physical layout of the Common Record XML Document file are required to ensure all systems transmitting or processing an instance of the file can do so without encountering physical constraint errors.

Examples of physical constraint errors include: exceeding available memory, exceeding available disk space, or exceeding available network bandwidth (and timing out). While most physical constraint errors can be avoided by proper coding techniques on the host system, there are some error conditions best handled through the definition of, and adherence to, a set of well-defined and well-publicized standards. Such error conditions and their corresponding standards for resolution are provided in this section.

XML and its associated technologies, (XML Schema, Document Type Definitions), do not provide for a mechanism to define the physical layout of a file. In fact, XML does not depend on the physical layout of the elements by design. Whitespace and line lengths are not relevant to the logical construction and reading of an XML document. Therefore, the Common Record XML file physical layout standards are defined as guidelines and examples presented in this Appendix.

32 K Line Length Limitation on Mainframe Systems

Issue:

Some mainframe systems cannot create or read a line longer than 32 kilobytes. A line consists of a string of data with an 'end of line marker.' The phrase 'end of line marker' in this document refers to whatever mechanism is used on a platform to terminate an individual line or record. For example, an end of line marker can be set by the writing out of a record (on a mainframe), by inserting a <CR><LF> character sequence (on a PC), or by inserting a <CR> character (on Unix). <CR>

(carriage return character) and <LF> (linefeed) are ASCII codes 10 and 13, respectively.

Systems with this limitation cannot handle a string of data longer than 32 kilobytes without encountering some type of data corruption (most likely truncation). By rule, XML ignores whitespace, which includes tabs, linefeeds, and carriage returns. Since XML has no means to regulate line lengths, a separate standard must be set to ensure lines do not exceed 32 kilobytes.

Solution:

The solution, in its general form, is to make sure an end of line marker is always set before a particular output stream of data reaches 32 kilobytes. This end of line marker should be created in the format native to the system where the file is being created.

The data transmission software used by COD and by participating schools handles the translation of end of line markers across platforms. For example, if a file were sent from a PC to a Unix box, the <CR><LF> character sequence would be automatically changed to <CR> characters. Therefore, the insertion of an end of line marker does not have to take into consideration the platform of the destination system. This same principle holds true between mainframes (EBCDIC or ASCII), PCs and UNIX boxes.

Note: The following standard is a strongly recommended approach for schools and vendors who want to guarantee their files will not exceed the 32 kilobyte limitation. If a school or vendor can ensure their submitted files will not exceed the 32 kilobyte limitation by some other means, those files will be accepted by COD. However, the burden of responsibility will then rest with the school or vendor for files incorrectly submitted.

Standard

In order to ensure no line grows larger than 32 kilobytes, all Common Record XML files submitted to COD should be constructed with end of line markers inserted at specified points in the document. These points are tied to specific elements in the XML document. The points are:
- After the Attending School element opening tag, - <AttendingSchl>
- After every Student element closing tag, - </Student>
- Before beginning an 'Award' element opening tag, which could be <Pell>, <DLSub>, <DLUnsub>, <DLPLUS>, <Perkins>, <SEOG>, <FWSP>, <CWC>, and the opening tags of all other elements designated as part of the Award substitution group in the Common Record XML Schema specification

- Before beginning a Disbursement element opening tag, -<Disbursement>, for every disbursement under each of the 'Award' elements listed above.

The Common Record XML Schema has been evaluated to confirm that if the above standard is maintained, the maximum size of the data sent in any one of these lines will never exceed 32 kilobytes. The maximum line size takes into account all tags, all maximum data lengths for each element, and all maximum number of repetitions of nested data elements.

The example found on the following page is a skeleton submission file correctly formatted according to the end of line marker standards. In this example, for clarity, sub-elements are not listed, and the characters, [EOL], explicitly mark the end of a line.

COD guarantees the Response Document files it produces will not contain lines longer than 32 kilobytes, but the files will not necessarily match the line breaks as specified for submissions. This is due in part to the additional space used by the Response blocks.

Coding Guidelines

The code should track the Common Record elements as they are written to the output stream and write out a line of XML text (via the appropriate end of line marker) as any of the EOL points are encountered. Code to construct the Common Record can vary greatly depending on the platform, libraries, and languages used, so specific examples are not provided.

```
<CommonRecord><!--All Common Record sub elements –
><ReportingSchl><AttendingSchl> [EOL]
<Student><!--All Student sub elements until an Award is encountered-->[EOL]
<Pell><!--All Pell sub elements until a Disbursement is encountered-->[EOL]
<Disbursement><!-- sub elements --></Disbursement>[EOL]
<Disbursement><!-- sub elements --></Disbursement>[EOL]
<Disbursement><!-- sub elements --></Disbursement></Pell>[EOL]
<DLSub><!--All DLSub sub elements until Disbursement is encountered-->[EOL]
<Disbursement><!-- sub elements --></Disbursement>[EOL]
<Disbursement><!-- sub elements --></Disbursement>[EOL]
<Disbursement><!-- sub elements --></Disbursement></ DLSub ></Student>[EOL]
</AttendingSchl><AttendingSchl>[EOL]
<Student><!--All Student sub elements until an Award is encountered-->[EOL]
<DLSub><!--All DLSub sub elements until Disbursement is encountered-->[EOL]
<Disbursement><!-- sub elements --></Disbursement></ DLSub ></Student>[EOL]
</AttendingSchl></ReportingSchl></CommonRecord>[EOL]
```